2017 National Geospatial Data Asset (NGDA) Dataset Report
USGS National Geologic Map Database Collection

Introduction:

This report aggregates information related to the National Geospatial Data Asset (NGDA) in the title above. In order to be designated an NGDA, the dataset must meet the criteria outlined in the Office of Management and Budget OMB Circular A-16 Supplemental Guidance. The guidance also directs Federal agencies to implement and use a portfolio management approach to ensure NGDA Datasets are managed by officially designated agencies, on behalf of all users, as national capital assets. As part of this process, the NGDA Dataset Managers regularly assess the maturity of their NGDA Datasets based on the geospatial data lifecycle and agency business requirements. All NGDA Datasets are assessed uniformly using a set of benchmark questions and a maturity index. This report includes results from the 2017 Lifecycle Maturity Assessment (LMA) which will be used to inform NGDA Dataset Managers about priorities and will be aggregated into a Theme Summary Report for NGDA Theme Leads.

NGDA Dataset Details:

Official Title: USGS National Geologic Map Database Collection

Metadata Record Title*: USGS National Geologic Map Database Collection

Theme: Geology

Dataset Lead Agency: U.S. Department of Interior, U.S. Geological Survey

Theme Executive Champion(s):

Name: Thomas (Tod) Dabolt
Agency: U.S. Department of Interior
Email: thomas_dabolt@ios.doi.gov

Theme Lead(s):

Name:	Dave Soller	Paul Godfriaux
Agency:	U.S. Geological Survey	Bureau of Ocean Energy Management
Email:	drsoller@usgs.gov	paul.godfriaux@boem.gov

Dataset Manager(s):

Name: Dave Soller
Agency: U.S. Geological Survey
Email: drsoller@usgs.gov

** If the metadata has been updated and re-harvested after publication of this report, the link may no longer be valid. The NGDA Dataset may be searched for manually in Data.gov or GeoPlatform.gov using the metadata record title.*

2017 Lifecycle Maturity Assessment (LMA)

Reporting on the status of each NGDA Dataset is an OMB requirement, and assessing the developmental maturity of the NGDA provides managers the ability to support NGDA Datasets in a more universal and transparent manner. The LMA is based on the OMB Circular A-16 seven stages of the geospatial data lifecycle with associated benchmark activities for each stage. Each benchmark has a range of activities, from no activity to significant activity, which provide the interpretation of maturity for that benchmark. The cumulative level of activity determines the maturity of the each NGDA Dataset and is based on a Maturity Matrix and How to Calculate Maturity. Additional information can be found at the GeoPlatform.gov 2017 Lifecycle Maturity Assessment (LMA) community web page.

2017 LMA Questions and Responses

The 2017 LMA includes a series of questions about the benchmark activities within each lifecycle stage, an explanation of specific actions that might be used to accomplish the benchmark activity (clarifying statements), and a series of status metrics in the form of response options. The NGDA Dataset Manager selects the response option that most accurately describes the current maturity level for the NGDA they manage. The responses to all the questions, and their associated metrics, collectively determine the overall maturity of an individual NGDA. The 2017 LMA consists of 3 general questions and 20 maturity questions. An additional 12 questions justify response choices to the 20 maturity questions. The questions are organized across the seven stages of the geospatial data lifecycle or pertain to all lifecycle stages. The detailed 2017 LMA questions can be found in the National Geospatial Dataset Asset 2017 Lifecycle Maturity Assessment document.

General Questions for All Stages

0) Part 1: Is this dataset considered "active" or "static"?
Response: The dataset is considered active - the dataset is being actively updated and maintained and has active components in some or all of the 7 Lifecycle Stages (with the exception of the Archival Stage)

0) Part 2: For the 2017 LMA, please provide what time frame the assessment includes?
Response: For an NGDA that completed the baseline assessment, responses include dataset activities from 2015 to 2017

0) Part 3: Do you have a process to evaluate the dataset to determine if it continues to meet the criteria established for a National Geospatial Data Asset (NGDA)?
Response: Yes

1) Is there a recurring process to obtain funding for all lifecycle stages of this dataset?
Response: Funding is currently adequate and consistent but tied to business requirements whose appropriations are not directed to support all lifecycle stages of the NGDA (see note above)

1.a) To justify your response to Question 1, what is the primary funding source for your NGDA?
Response: Directed appropriation(s) - Funding signed into law by Congress for a specific program that supports this NGDA or the NGDA itself

1.b) To further justify your response to Question 1, do you have secondary source(s) of funding?
Response: Yes

1.b.a) What are the other source(s) of funding for your NGDA?
Federal interagency - Agreements between two or more different Federal Department(s), Agency(ies), Bureau(s), and/or Office(s) to provide funding for a specific program that supports this NGDA or the NGDA itself
Other - The dataset for geologic mapping is a collection of data, observations, and interpretations that are found in individual geologic maps and reports, both published and unpublished. This vast collection is supported by an evolving set of geoscience and technical standards, developed in cooperation among the partners who contribute to the collection. Together, this standardized collection ("dataset" in NGDA terms) is designated as the National Geologic Map Database (NGMDB, http://ngmdb.usgs.gov/). This collection includes some individual GIS datasets, but the majority of the information is in the form of scanned (raster) copies of individual maps and reports, and the bibliographic citations and web links to downloadable copies. The NGMDB is Congressionally mandated, in the Geologic Mapping Act of 1992 (http://ncgmp.usgs.gov/about/ngm_act/ngmact1992.html) and its Reauthorizations (most recently in 2009, http://ncgmp.usgs.gov/about/ngm_act/ngmact2009.html). The Geologic Mapping Act established the National Cooperative Geologic Mapping Program (NCGMP, http://ncgmp.usgs.gov), administered by the U.S. Geological Survey (USGS) in order to provide funding for geologic mapping by USGS, the State Geological Surveys, and Universities. The NGMDB is identified as the Act's sole purpose, in Section 31a(b) as follows: "The purpose of sections 31a to 31h of this title is to expedite the production of a geologic-map data base for the Nation, to be located within the United States Geological Survey, which can be applied to land-use management, assessment, and utilization, conservation of natural resources, groundwater management, and environmental protection management." The Geologic Mapping Act of 1992 established the NCGMP as the funding source for geologic mapping and for the NGMDB. When the NGMDB was initiated in 1995, funding was adequate to support all requirements. Over the succeeding 20 years, funding decreased to approximately 20-25% of the 1995 level. Concurrently, computer technology and the public's expectations for delivery of information have advanced dramatically, necessitating a significant broadening of the NGMDB's scope and increased investment in technology. Therefore, funding support is quite inadequate to support public and legislative demands on the NGMDB.

2) Is there a process in place to ensure that open government and transparency guidelines are followed in all lifecycle stages for this dataset?
Response: Implementation well established

3) Are there processes and tools in place so that staff are sufficiently knowledgeable to ensure the continuity of the dataset for all stages of the lifecycle, especially during staffing transition?
Response: Implementation progressing

STAGE 1 - Define/Plan

Characterization of data requirements based upon business-driven user needs.

4) Are business requirements defined and formalized?
Response: Fully implemented including recurring assessments

5) Are there processes in place to ensure partners and stakeholders are involved in requirements collection?
Response: Fully implemented including recurring assessments

5.a) To justify your response to Question 5, which external partners and stakeholders are involved in the requirements collection process?

Federal agency - Partners and stakeholders are participants from other federal agencies
State government - Partners and stakeholders are participants from state governments
Professional association - Partners and stakeholders are participants from professional associations
Other - The Geologic Mapping Act of 1992 stipulated the NGMDB as a collaborative responsibility of the USGS and the State Geological Surveys (represented by AASG, the Association of American State Geologists). Therefore, the NGMDB was defined and established after extensive discussion amongst these stakeholders, culminating in 1995 in a public statement of intent (http://ngmdb.usgs.gov/Info/reports/geotimes95.html). Continuously since then, the NGMDB has been reexamined by the stakeholders and refined as needed. This reexamination is facilitated by discussion at public venues (e.g., scientific and stakeholder meetings, and the annual Digital Mapping Techniques workshops (http://ngmdb.usgs.gov/Info/dmt/)) and in published techniques and standards 5) How are partners/stakeholders involved in the requirements collection process? Answer: Justification Comment: STAGE 1 - Define/Plan Attachment(s): 0 Attachment(s): 0 Attachment(s): 0 Created: 2015/12/31 NGDA Dataset Report | 5 articles and reports of progress (e.g., in Proceedings of the Digital Mapping Techniques workshops). Other stakeholders have been identified (e.g., the National Park Service, the Association of Independent Professional Geologists, American Association of Petroleum Geologists), and their input and requirements are carefully considered.

6) Is there a quality assurance process for the dataset?

Response: Implementation well established

6.a) To justify your response to Question 6, what methods are used to develop and complete quality assurance assessments?

Internal quality assurance assessment - A defined review process is undertaken by agency personnel through quality evaluation, testing, or other manual or automated methods
Other - Numerous quality-control measures have been developed, and are implemented continuously. These include: (1) scripts to check for broken links to collaborator and other publisher's web sites and downloadable maps and reports; (2) programming logic to identify errors in data input (e.g., bibliographic citations and links) by collaborators; (3) daily use of the NGMDB by project staff, through which content errors are identified and corrected; and (4) errors reported by users, via the Corrections web form.

7) Has an assessment been done to evaluate the sensitivity, privacy, and confidentiality of this dataset?

Response: Fully implemented including recurring assessments

8) Are defined data standards used in collecting, processing, and/or rendering the data?

Response: Implementation progressing

8.a) To justify your response to Question 8, what types of data standards are used in collection, processing, and/or rendering the data?

FGDC-endorsed standards in use: Geologic Map Symbolization (http://ngmdb.usgs.gov/fgdc_gds/). The following is information from 2015 report: Standards for geologic mapping, including geologic time and stratigraphic nomenclature, have existed since the late 19th century. These range from agency-adopted common practices to widely accepted standards developed by international bodies (e.g., the North American Commission on Stratigraphic Nomenclature). A common suite of standards are generally followed by the Nation's principal agencies that conduct geologic mapping (i.e., USGS and the State Geological

Surveys). Differences in geologic map content and emphasis (e.g., mineral or water resources, or hazards) do exist, but the fundamental science aspects (e.g., geologic time, conceptual agreement on what constitutes a mappable geologic unit, stratigraphy, lithologic descriptors), are reflected in the relatively uniform nature of geologic maps. The NGMDB is charged under the Geologic Mapping Act of 1992 with reexamining these standards and updating them as needed for implementation in a digital publication process and incorporation in the Database. The cartographic standards used for conventionally printed maps have been updated for digital maps, and released as a FGDC Standard (http://ngmdb.usgs.gov/fgdc_gds/). A format and content standard for publication of the geologic map database (http://ngmdb.usgs.gov/Info/standards/NCGMP09/) is nearing final revision before being proposed as a FGDC Standard. Other standards, including NGDMB-specific workflow and content, are completed or under development (see various links under http://ngmdb.usgs.gov/Info/).
Standards not endorsed by FGDC in use: NCGMP09 (GeMS)
http://ngmdb.usgs.gov/Info/standards/NCGMP09/

STAGE 2 - Inventory/Evaluate

The creation and publication of a detailed list of data assets and data gaps (both internal and external) as they relate to business-driven user need.

9) Is an assessment done to determine if data necessary to meet requirements already exists from other sources (either within or outside the agency) before collecting or acquiring new data?
Response: Fully implemented including recurring assessments

9.a) To justify your response to Question 9, what actions are performed to determine if data already exists from other sources?

Internal research and evaluation - The lead federal agency conducts research to identify whether other potential sources and datasets exist, including GeoPlatform.gov Marketplace research
External research and evaluation - External stakeholders and partners conduct research to identify whether other potential sources and datasets exist
Suitability review of datasets - Datasets identified that may meet mission requirements undergo suitability review or testing
Financial cost/impacts evaluation - Datasets identified that may meet mission requirements undergo review for potential financial costs or savings
Other - Geologic mapping is conducted almost exclusively by State Geological Surveys and by a few Federal agencies. These geologic mapping organizations are in close contact with one another, which minimizes the possibility that suitable, preexisting mapping already exists for an area that is proposed to be geologically mapped. The NGMDB Map Catalog also serves as the authoritative record of all geologic mapping of the Nation, and can be consulted in order to identify potential duplication of effort. Most significantly, agency review boards (e.g., State Mapping Advisory Boards) determine the areas to be geologically mapped; this serves to prioritize areas for geologic mapping, and strongly considers any possibility that previous mapping of the area is suitable for the societal issues that need to be addressed.

STAGE 3 - Obtain

The collection, purchase, conversion, transformation, sharing, exchanging, or creation of geospatial data that were selected to meet the business needs is identified.

10) Is there a process for obtaining data for this dataset?
Response: Fully implemented including recurring assessments

10.a) To justify your response to Question 10, what actions are performed to obtain data?

Obtain data by sharing/exchanging

Obtain data by creating and/or collecting the data

Other - The National Cooperative Geologic Mapping Program (NCGMP, http://ncgmp.usgs.gov) administers the funding of a large proportion of the Nation's geologic mapping. Through the NCGMP's process of identifying areas to be geologically mapped, new information is added to this dataset (the NGMDB) on a continuous basis. There are other Federal and State programs that fund geologic mapping; the maps from these programs are incorporated into the NGMDB to the extent possible (i.e., given resources available to identify these maps and add bibliographic content to the NGMDB).

11) Is the metadata in an FGDC-endorsed geospatial metadata standard, follows the NGDA Metadata Guidelines, and is published?

Response: Published and implementation of NGDA Metadata Guidelines well established

12) Part 1: Is there a business process in place to determine the geographic coverage of the dataset and establish milestones to track progress towards completion?

Response: Fully implemented including recurring assessments

12) Part 2: Based on the business requirements, what is the estimated completeness of the geographic coverage?

Response: Geographic coverage is more than 50% but less than 75% complete based on business requirements

12.a) To justify your response to Question 12 Part 2, what is the geographic coverage of the dataset as defined by the business requirements?

Response: Other

Response: The NGMDB includes bibliographic citations for >101,000 geoscience maps and reports; among these are >38,000 publications focusing on bedrock or surficial geologic mapping. The coordinates of each 12) How complete is the geographic coverage as defined in the requirements for the dataset? Part 1 Answer: Justification Comment: Data set is roughly 50% of the geographic coverage is presently complete per current requirement. Part 2 Answer: STAGE 2 - Inventory/Evaluate STAGE 3 - Obtain Attachment(s): 0 Attachment(s): 0 Attachment(s): 0 Attachment(s): 0 Answer: Created: 2015/12/31 NGDA Dataset Report | 7 map's bounding box are recorded in the NGMDB. However, many geologic maps are not rectilinear in outline (e.g., a geologic map of a mining district, or a particular belt of rocks), and so a computation of the areal coverage of such geologic maps would be highly inaccurate. Therefore, in response to government requirements for "productivity metrics" for funding provided by the National Cooperative Geologic Mapping Program, the NGMDB project computes the geographic coverage of geologic mapping by considering only those maps that are rectilinear, or presumed to be nearly so. The process of determining which maps to include in the computation is complicated and relies on judgment and experience in performing this task each year, in order to be as consistent as possible from year to year. The basic methodology is described in http://pubs.usgs.gov/of/2005/1428/pdf/soller2.pdf. Because the process summarizes the total land area for which a geologic map is available, it is largely dependent on state and locally prioritized needs for mapping, which focus on the most pressing societal issues. If these societal issues applied equally to all areas of the Nation, then it could be foreseen that all areas of the country would soon be mapped at the relatively detailed scales (1:100,000 and more detailed) that are considered by this metric. However, for vast areas of the Nation (e.g., in Alaska and the midcontinent), geologic mapping has been done only at a reconnaissance level (e.g.,

1:250,000), and it is unlikely given the nature of the geology, the population density in those areas, the nature of the most pressing societal issues, and the level of funding, that a large amount of more detailed mapping will be conducted in the near future. Instead, NCGMP funding more commonly is applied to remapping high-priority areas, with the purpose of providing more detailed and modern geoscience information. With these caveats, we can state that, at scales of 1:100,000 and more detailed, 53.5% of the Nation has been geologically mapped.

STAGE 4 - Access

Making data produced known and retrievable to the community through documentation and discovery mechanisms so the users can meet their business requirements.

13) Do you provide users access to the data in a digital machine-readable format?
Response: Fully implemented including recurring assessments

13.a) To justify your response to Question 13, what types of digital machine-readable formatted web services or data download services are available for this dataset?

KMZ/KML - Keyhole Markup Language
SHP - Shapefile
ASCII - American Standard Code for Information Interchange
Esri Grid format
Geospatial TIFF - Geographic Tagged Image File Format
GeoPDF - Geospatially enabled Portable WORD Format
DEM - Digital Elevation Models
MS - Excel with XYZ values
GDB - File Geodatabase
MDB - Personal Geodatabase
ArcInfo Coverage

STAGE 5 - Maintain

The ongoing processes and procedures to ensure that the data meet business requirements.

14) Is there a maintenance process in place for this dataset?
Response: Fully implemented including recurring assessments

15) Is there a quality assurance/quality control (QA/QC) process as part of this dataset's maintenance?
Response: Fully implemented including recurring assessments

STAGE 6 - Use/Evaluate

The ongoing assessment, validation, and potential enhancement of data to meet user needs and business requirements.

16) Is there a process to determine if the dataset meets user needs?
Response: Fully implemented including recurring assessments

17) Is there a process to provide users information on how to evaluate and properly use the dataset?
Response: Fully implemented including recurring assessments

18) Do the business processes and management practices include an assessment of changing technology?
Response: Fully implemented including recurring assessments

STAGE 7 - Archive

Facilitate the selection/appraisal retention, storage, preservation and accessibility of geospatial content with long-term value (or the disposition of material as appropriate) and establish mechanisms for the coordinated development of stewardship tools and services across all impacted Federal agencies.

19) Is there an archiving appraisal process for the dataset?
Response: Implementation well established

19.a) To justify your response to Question 19, where is digital data being archived as determined by your appraisal process outcome?

Other - The NGMDB holds ~25 TB of image content, workflow scripts, and an extensive RDBS of geoscience content that has been built over a 22 year period. At present, this information resides at the Denver Federal Center (DFC) on USGS servers and mass storage devices; the NGMDB project pays another USGS group for the use of these devices and for IT support. The system is backed up to a tape archive, and tapes are stored offsite. In order to provide redundancy and to increase the security of the data, NGMDB-owned servers that previously were use to serve data to the public are being moved to Reston, VA, as a backup system; they also will provide a staging area for processing new data. Because the NGMDB is a collaborative project with the State Geological Surveys, there is some built-in redundancy; some States maintain their bibliographic data in the identical form used by the NGMDB. Further, many of the map images used in the NGMDB are available through the State Geological Surveys.

19.b) To justify your response to Question 19, where is printed data being archived as determined by your appraisal process outcome?

Other - Regarding the NGMDB-managed unpublished source information (mostly paleontologic and stratigraphic reports) referred to in Question 2, these exist as paper records. They are curated in a climate-controlled facility in Reston, VA., along with thousands of older published documents, reports, and maps that are used daily by the NGMDB project in order to improve the content of the NGMDB dataset. Many of those documents are not available in the USGS Reston Library or on the Internet, and so serve as an invaluable resource to the project and to visiting scientists.

LMA Submission and Reviewer Information

LMA Submission:

Status: Complete
Date: 8/17/2017

Submitted by:

Name: Dave Soller
Email: drsoller@usgs.gov

Reviewed by:

Name: Jen Carlino
Email: jcarlino@usgs.gov
Date: 8/16/2017

LMA Maturity Overview

Each of the 20 maturity question responses was assigned a maturity level based on a Maturity Matrix. The maximum level of maturity (Optimized; Established) is level (5) and the least level of maturity (No Activity) is level (0). The question response maturity was then averaged across each lifecycle stage to determine the stage maturity as described in How to Calculate Maturity. The table below shows the numerical maturity levels with their corresponding descriptions and characteristics.

Maturity	Maturity Characteristics for All Lifecycle Stages
Optimized; Established Level = 5	Dataset meets virtually all business needs of all users. The dataset is considered authoritative by owners and secondary users. It is curated across all stages of the approved lifecycle. Future needs are defined for both the primary owner and secondary users on a regular basis and resources for addressing both current and future business requirements are available.
Mature; Consistent Level = 4	Dataset meets all the business needs of the primary owner and most of the secondary users. The dataset is curated and used as an authoritative resource by the primary owner and secondary users. Future needs are being identified and steps are planned to address these. All stages are supported and reviewed on a recurring basis. The dataset is well managed in relation to the approved lifecycle.
Managed; Predictable Level = 3	Dataset meets a significant number of the business needs of the primary owner and is widely used by secondary users. Benchmark activities are occurring in at least four of the approved lifecycle stages. Management practices in relation to the approved lifecycle is moderate but consistent. Dataset is integrating changing business requirements in lifecycle stages impacting overall maturity.
Transition; Transformation Level = 2	Dataset meets business needs of the primary owner and has moderate use by secondary users. Benchmark activities are occurring in at least three stages. Efforts to integrate funding, include partners, and obtain data are not supported in a sustained manner. Management practices in relation to the stages of the approved lifecycle is limited.
Planned; Initial Development Level = 1	Dataset in initial planning or limited in meeting business needs of the primary owner. Benchmark activities in the approved lifecycle are just starting to consider secondary uses, Partners/stakeholders involvement is being defined and developed to support additional dataset uses. Dataset development is in a very early stage. Minimal or limited management against the benchmarks in the approved lifecycle.
No Activity Level = 0	Dataset not developed or meets project/local business needs of the primary owner. Secondary, additional uses, or partners/stakeholders were not considered. Dataset is not recognized as authoritative data or is part of a similar dataset. Not managed to any of the benchmarks in the approved lifecycle.

Table 1: 2017 Maturity Matrix.

2017 NGDA Dataset Maturity Results

Based on the maturity question responses, an overall maturity level was calculated for this NGDA, along with maturity calculations for the general questions for all stages and each discrete lifecycle stage as shown in Table 2 below.

Maturity Categories	Maturity Level
Overall NGDA Maturity	**4 - Mature; Consistent**
General Questions for All Stages	4 - Mature; Consistent
Stage 1: Plan/Define	4 - Mature; Consistent
Stage 2: Inventory/Evaluate	5 - Optimized; Established
Stage 3: Obtain	4 - Mature; Consistent
Stage 4: Access	5 - Optimized; Established
Stage 5: Maintain	5 - Optimized; Established
Stage 6: Use/Evaluate	5 - Optimized; Established
Stage 7: Archive	4 - Mature; Consistent

Table 2: 2017 Maturity Results.

LMA Process Changes Between 2015 and 2017

In 2015, a baseline assessment of National Geospatial Data Assets (NGDA) was performed for each of the NGDA Datasets in the federal geospatial portfolio. Information related to the 2015 baseline LMA can be found at 2015 NGDA Lifecycle Maturity Assessment, which also includes a link to the 2015 Reports. A follow up analysis of the 2015 LMA baseline process and its results identified ways to improve the LMA workflow, increase efficiency as well as decrease reporting burden. Several recommendations were identified and implemented in 2017, which included improvements to normalize the responses in 2017. A secondary effect of improvements to normalization is that results from 2017 and 2015 are not directly comparable. These changes, and their impacts, are detailed in the webpage: Temporal Changes in Lifecycle Maturity Assessment (LMA) Maturity and Results Comparisons.